2009

Experimentation and Modeling of Soldier Target Search

Chung, Timothy H.

Matthew Hasting, Timothy H. Chung, and Paul Evangelista. "Experimentation and Modeling of Soldier Target Search." Proceedings of the International Conference on Industrial Engineering,

Proceedings of the 14th Annual
International Conference on Industrial Engineering
Theory, Applications and Practice
Anaheim, California
October 18-21, 2009

EXPERIMENTATION AND MODELING OF SOLDIER TARGET SEARCH

Matthew Hasting[1], Timothy H. Chung[1], and Paul Evangelista[2]

[1]Department of Operations Research
Naval Postgraduate School
Monterey, California 93943

[2]Training and Doctrine Command Analysis Center – Monterey
Naval Postgraduate School
Monterey, California 93943

Corresponding author's e-mail: matthew.hasting@us.army.mil, thchung@nps.edu, paul.evangelista@us.army.mil

Abstract: This paper investigates the visual search process and the effect of contextual information on the search process in an urban combat environment. High resolution combat simulation models implement a parallel sweeping or "windshield wiper" search process that is not representative of human search behavior. Furthermore, combat models do not account for additional situational awareness in the form of contextual information. A discrete myopic search model is proposed, a study of which provides a statistical model based on human performance data. This model prioritizes search effort where humans believe that targets are most likely to occur. Nineteen volunteers searched 16 static urban scenes with zero to five targets. These data formed the probabilities that a target is located in each cell in each discretized scene. The discrete myopic search model chooses the cell with the highest probability for each discrete look. Hypothesis testing on experimental data revealed a nearly 20% increase in accurately predicting human search patterns using the discrete myopic search model over the windshield wiper model. Further investigation revealed a significant change in search behavior and detection performance based on the addition of contextual information. The major result of this work indicates that combat models need to bias search based on the situational awareness of observer and properties of the observer's environment.

1. INTRODUCTION

Search and target acquisition (STA) in military models and simulations attempts to abstract arguably the most important aspects of combat operations. The process of a combatant observing another combatant, deciphering the combatant's propensity for hostility, and then deciding whether or not to engage the combatant with lethal force nearly summarizes individual close order combat. The methods developed to represent this process will alter the outcome and validity of a combat model. Modeling STA that occurs with the assistance of a technical device such as a scope or radar significantly simplifies the modeling problem. Measured performance parameters of the technical device bound the STA process, and computer models can replicate these quantifiable properties. Representing human STA with the unaided eye does not offer such a simple solution. Human vision does not have easily bounded parameters. Although certain components of human vision are well understood, other components remain a mystery. Scientists understand that humans perceive a fraction of the electromagnetic spectrum, extract meaning from color and shape, and see clearly at a two degree foveal point but clarity degrades in the periphery. Other human vision phenomena are not entirely codified. Consider the desire of a person to fixate at certain points within a scene but ignore others, anticipate threat or danger from particular areas within view but quickly surmise other areas as impotent, or interpret defined detection goals from an environment based on previous information or activity. Psychological theories detail the likely cognitive processes that lead to these types of behaviors, but predicting human visual behavior given complex and natural visual stimuli remains an open problem (Doll and Home, 1999, Jungkunz, 2009). Bruce et al. (2005) summarize visual search as the product of two factors: the properties of the surrounding environment and the goals of the observer.

A major flaw that currently exists in combat models involves the application of unbiased search. The underlying STA algorithm accepted and widely used in many combat models can be traced to an experiment performed by Johnson that essentially correlated and summarized the difference between target contrast and detection by the observer (Johnson, 1958). This experiment, supported largely by the U.S. Army Night Vision and Electronic Sensors Directorate (NVESD), was developed to support modeling of aided eye detection with a restricted field of view, such as looking through a scope. The current detection algorithm in most U.S. military combat models, known as ACQUIRE, implements search patterns based upon the search strategy recommended for the vision enhancing equipment that will be modeled. Although there are recommended search strategies for the unaided eye, research supports that pre-attentive processing of environmental surroundings supported by human learning, specifically our ability to aggregate and associate to build context, trumps any attempts to train or control human search behavior (Medina, 2008, Hoffman, 1998).

The work presented in this paper serves to motivate the problem of modeling unaided human search, specifically applied to the task of target acquisition in combat operations. The primary contributions of this paper include (i) results from an experiment conducted to explore search in combat environments, and (ii) discussion and results from a model that refutes unbiased search. An important aspect to remember throughout this paper involves the task of model integration. Impressive research has been performed in the area of STA, described in an excellent survey by (Vaughn, 2006). Many of the models surveyed by Vaughn develop what appears to be a promising approach to improve the fidelity of STA representation, however the careful reader will notice shortcomings. Most of these shortcomings will arise during implementation, and Vaughn points out many of these. Examples of these shortcomings include a requirement to include parameters or variables that are unreasonable to expect from current combat models, a requirement for free variables that preclude the ability to generalize in closed form combat models, or computational complexity that creates an unacceptable processing burden. These advanced approaches need to be pursued in the hope and strong likelihood that combat models will continue to progress and soon be able to accommodate this level of complexity. However, it is also important that research and development efforts also focus on STA models that can make a near term impact with existing combat models. In order for an STA model to make a near term impact, strong potential for improvement of STA representation must be evident and the modeling community must be willing to accept the representation. The former requirement manifests from experimentation and analysis; however, the latter requirement means that the proposed solution should be implemented within the existing model architecture and should not create an unreasonable processing burden. If these requirements are not met, even the best proposed solutions are unlikely to make a difference in the near term.

2. AN EXPERIMENT TO MEASURE SEARCH BEHAVIOR IN TARGET ACQUISITION

An experiment designed to measure target search in a combat environment took place at the Naval Postgraduate School in December of 2008. The experiment had a singular objective: determine the effect of situational awareness (SA) on STA. Situational awareness is a field of research in and of itself (Endsley, 2000). Miller and Shattuck describe the effect of SA on decision making as a "dynamic model of situated cognition". This model portrays data available as only a subset of ground truth, morphing data through lenses that synthesize and associate information based on individual experience and local context, eventually becoming information used to support a decision (Miller and Shattuck, 2006).

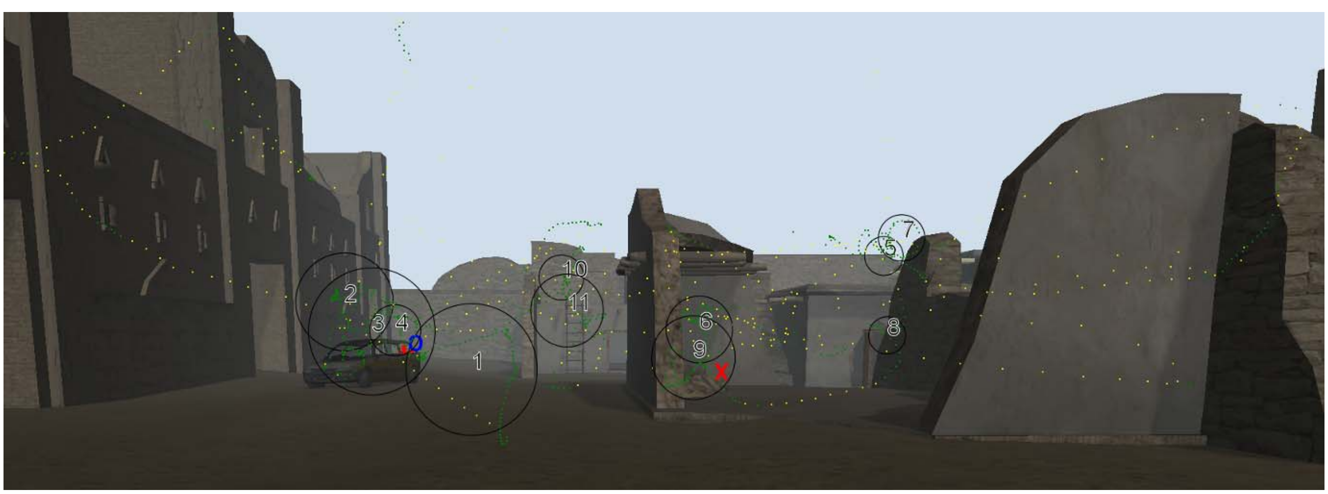

Figure 1. This cropped example scene has eye-tracking results overlayed. This particular scene contained an audible report referencing the car and a moving target behind the car. Yellow dots represent high velocity eye movements (>12.5° per second), green dots represent low velocity eye movement (<12.5° per second), and circles represent long fixations, areas of sustained low velocity eye movement for 20 µs or longer. Long fixations are numbered in chronological order.

The experiment consisted of sixteen scenes presented to nineteen participants, and the design of experiments (DOE) included four binary factors: a moving target, an audible report, a written situation report (SITREP), and a minimap. Each scene consisted of an urban environment with a unique field of view and array of structures. The moving target varies from 0.2 – 0.4 degrees in size, and it moved into the scene for approximately one second and then moved back out of the scene. The audible report consisted of a brief (5-10 seconds) report indicating a target location with a semantic reference. The written situation report was a sheet of paper that provided several bulleted remarks about the physical nature of the environment and likely enemy courses of action along with an overhead schematic of the scene. Each scene lasted 20 seconds, and the participants indicated detection of a target with a mouse click on the target location. Eye tracking was recorded using the Seeing Machines FaceLab4 eye tracker. Participants sat 71 cm from a 24 inch TFT monitor set to 60 Hz at a resolution of 1920 by 1200 pixels. Only results from participants achieving an eye tracking screen calibration error of one degree or better were included in the analysis. An example scene along with eye tracking results from a single participant is shown in figure 1. Notice the strong center bias, presumably a result of the audible reference and movement, two basic aspects of SA included as factors in this experiment. It is also interesting to note the lack of interest in the outer portions of the scene. Minimal ambiguity and close range contribute to a person's ability to pre-attentively clear portions of scenes, resulting in the appearance of ignoring or paying minimal attention to a portion of the scene. Unambiguous, uncluttered, close range components of scenes received very little attention during this experiment. Assuming time fixated represents areas of perceived threat, the perceived threat resides in ambiguous, cluttered areas and areas where external information indicated immediate threat.

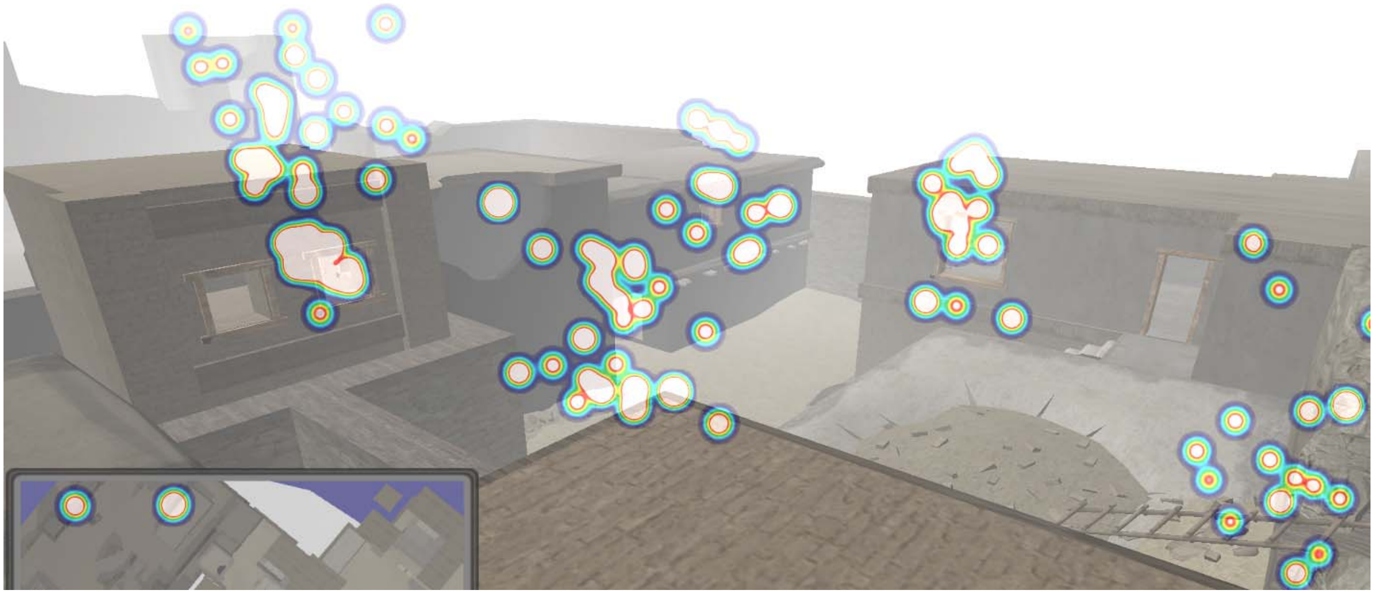

Figure 2. This cropped example scene has been overlaid with a heat map that reflects fixation duration aggregated over all the participants. Every fixation from every participant that observed this scene is shown, and a larger white heat mark indicates a longer fixation. Clusters of fixations reinforces that search is not random.

Consider the aggregated fixations shown in figure 2. This figure reinforces the concept that search is not random. Clusters of fixations clearly form around windows, edges of buildings, and cluttered areas. Notice that there are entire rooftop areas and uncluttered windows and doors where no fixations occurred, suggesting that pre-attentive processing cleared these areas. Jungkunz reinforces this concept in his thesis with an interesting single target experiment that systematically varies a hiding location, clutter, target saliency, and target eccentricity (Jungkunz, 2009).

3. APPLICATION OF A DISCRETE MYOPIC SEARCH MODEL

Discrete myopic search is a search technique that optimizes the probability of detecting a target for each of *k* discrete observations. The given search area is discretized into *n* cells and a probability that a target is within the i^{th} cell, denoted p_i, is assigned or derived from prior information. Each successive look in cell *i* has a probability q_i of not detecting the target and this probability is assumed independent of all other looks. Given k_i looks in the i^{th} cell, the model optimizes the probability of detection $p = \sum_{i=1}^{n} p_i(1-q_i^{k_i})$ subject to $k_1 \geq 0$ and a constraint on $\sum_{i=1}^{n} k_i \leq K; k_i \in (0,1,...,K)$ for *n* cells (Washburn, 2002). The model implements a greedy algorithm in which the *k* + 1st look is allocated to the cell with the greatest increase to the overall probability of detection. The increase in probability of detection in the i^{th} cell is shown in equation 1:

$$p_i(1-q_i^{k_i+1}) - p_i(1-q_i^{k_i}) = p_i(1-q_i)q_i^{k_i} \tag{1}$$

This implementation produces an optimal allocation of discrete looks because $p_i(1-q_i^{k_i})$ is a concave function of k_i (Washburn, 2002). A critical assumption to ensure concavity and the optimality of the resulting search pattern is the absence of false positive detections. Given the context in which false alarms, such as engaging a civilian, would incur a prohibitively high penalty, the assumption of no false positive detection errors is considered reasonable. Further research, either psychophysical or otherwise, is left for future study. Common applications of a myopic search model include anti-submarine warfare and Coast Guard search and rescue. A historic example is the search for the sunken USS Scorpion in 1968. Searchers constructed a prior probability map as a composite of nine separate priors. Searchers developed the nine priors based on different scenarios explaining how and where the sinking occurred (Wagner et. al, 1999). This thesis proposes a novel application of this model in the form of visual search.

This study slightly modifies the model to accommodate zero to six targets. Normalization of the prior probability map ensures that the sum of probabilities in all cells equals one. If detection occurs, the cell or cells in which the detected target resides are eliminated from future consideration. This change allows the model to find the optimal search pattern with multiple targets.

Discretization of each scene into appropriately sized cells is the first step in implementing the Discrete Myopic Search model. Discretization allows for investigation of each participant's fixations and provides some robustness to error in the eye-tracking hardware. The size of each cell is determined by the average fixation size across all participants for each scene. This number of pixels in each cell is $\overline{fix}^2$ where $\overline{fix} = \sum_{i=1}^{n} fix\_size_i / n$ for *n* fixations in a scene across all participants. Fixation size is determined by the time duration of the fixation but translates to a number of pixels, each pixel representing 16μs of low velocity eye movement. For example, every green pixel in figure 1 represents 16μs of low velocity eye movement. Each scene has an average fixation size of 76 pixels, resulting in 375 square cells. The objective of this research is to develop a model to predict an individual's search pattern. The hypothesis is that an individual will search in a manner in which to optimize his ability to detect targets during each successive fixation. He prioritizes his fixation pattern based on where he believes targets will most likely be located. This also corresponds to proposed schemes, such as saliency-based methods for visual attention in which the scan pattern is prioritized based on decreasing saliency (Itti and Koch, 2001). Since participant fixations were used to determine the prior probability map, it is necessary to prioritize the fixations using weights.

Search performance is not constant even over a short period of time (Cooke, 1983). This rate of change is not well defined, but can be estimated from the experimental data. As a participant locates targets or fails to locate targets, his belief that targets remain in the scene decreases.

4. EXPERIMENTAL RESULTS

There is one *p* value and one *q* value for each cell in the discretized image resulting in two matrices, **P** and **Q**. Thirteen of the nineteen participant data sets were used to build **P** and **Q** and data from six participants was reserved to test the model - essentially measure the ability of the model to generalize. A comparison between the discrete myopic search model and windshield wiper search should reveal whether or not search is an unbiased activity. Strong performance of the discrete myopic search model indicates similar search patterns across a sample of participants. If stimuli and information biases search patterns, similar search behaviors will result and the discrete myopic search model implementation proposed in this paper will outperform windshield wiper search, in terms of better predicting true human visual search behavior.

The probability of a target in cell *i*, p_i, is based upon data taken from fixations of half of the subjects. Fixation duration in each cell, aggregated across all participants, generates p_i . The order of fixations is also included in this model. Analysis of detection statistics indicated that most of the detections occurred in the first few seconds of scene presentation, and it is important to include this result in the construction of **P**. Analysis of the data revealed that target detection over time followed a positive skewed gamma distribution, indicating that very early fixations were unsuccessful, quickly followed by spike in successful fixations, then degrading exponentially over time. Weighting fixations based on when they occurred in the scene resulted in higher p_i values in the cells where early fixations tended to occur. Normalization of the weighted aggregate fixations results in p_i summing to unity across all cells for a given scene. Detection statistics build the probabilities assigned to q_i. The probability of not detecting the target given a look in cell *i*, q_i, is derived from detection data taken from the same subjects. q_i is the probability that a target was not detected in cell *i* given a fixation in cell *i*.

The focus of this research is the modeling of the search process and not detection performance. It is paramount to this study that the detection probabilities are consistent for each model so that detection order and not detection performance forms the basis for comparison. The training data from thirteen participants provides the inputs implemented by the models below. The desired outcome is not to predict the probability of detection for the test set, but to instead predict the participants' search patterns by examining the order in which participants locate targets. Given that this is the goal of this research, measures of effectiveness (MOEs) become a bit more complicated. Two MOEs are considered, and both focus on target detection order. The first MOE (MOE1) is the probability of detection order given that a detection occurs. It is defined in formula 2 as the probability of detecting target *j* in order *k* given a detection in the scene. Let $d_{i,j,(k)}$ be a binary variable that equates to one if participant *i* detects target *j* in order *k* for a given scene, else it equates to zero, and let *T* equal the total number of targets in a scene. MOE1 is defined in equation 2.

$$P(d_{i,j,(k)} = 1 \mid \text{detection}) = \frac{\sum_{i=1}^{13} d_{i,j,(k)}}{\sum_{k=1}^{T} \sum_{j=1}^{T} \sum_{i=1}^{13} d_{i,j,(k)}} \tag{2}$$

MOE2 (equation 3) is the absolute probability of detection order for each target. It is the conditional probability of detecting target *j* in order *k* given that target *j* is detected.

$$P\left(d_{i,j,(k)} = 1 \mid \sum_{k=1}^{T} d_{i,j,(k)} = 1\right) = \frac{\sum_{i=1}^{13} d_{i,j,(k)}}{\sum_{k=1}^{T} \sum_{i=1}^{13} d_{i,j,(k)}} \tag{3}$$

These conditional probabilities could each be considered the probability of success in a binomial process. Given this, it is possible to build models from the training data and test model performance against the actual human performance in the test data. Both discrete myopic search and windshield wiper search were tested. Neither model performed exceptionally well, however there was a considerably better performance observed from the discrete myopic search model. Using the binomial hypothesis test at an alpha of 0.05, windshield wiper search was rejected as an equivalent binomial process 85% of the time compared to the discrete myopic search model which was rejected 70% of the time. Given that this percentage resulted from 1000 iterations of both windshield search and the discrete myopic search model, this difference is significant. With regard to the four factors included in this experiment, the audible report created the most significant effect. Over 93% of the first five fixations in scenes with an audible report were directed at the target indicated by the report. Although study continues with this data, the other three factors do no appear to create a significant change in search behavior.

5. CONCLUSION

Modeling human vision is a challenge. It is widely accepted that information, goals, experience, and environmental conditions affect STA. Quantitatively describing how these variables affect search and developing a concise method to implement this behavior in combat models remains as open research. (Doll and Home, 1999) reinforce many of these challenges and present recommendations for future research that remain relevant ten years later. Doll and Home clearly state that windshield wiper search is simply a poor assumption. Their discussion reinforces that scene structure (edges, corners, clutter) attracts attention and temporal sequence, the same concept explored by the MOEs in this paper, matters. However, ten years later windshield wiper search is still the dominant search method used in combat models.

The discrete myopic search model presented in this paper illuminates several aspects with regard to search. The most significant result that emerges from the analysis of this model is that search is not random. This model indicates that underlying patterns within scenes drive search behavior. Uncovering these patterns in a concise, generalizing form remains a challenge. Furthermore, quantifying the effect of information also remains as an open research topic. Simple, overt cues, such as gunfire or disruptive noise, have long been known to affect attention. The audible reports in this study take a small step forward in cognitive analysis, requiring some level of interpretation from the subjects to apply the information. It is the opinion of these authors that a holistic yet simple approach is necessary to improve search behavior in combat models. This holistic approach will likely rise from simple roots, such as the two factors that (Bruce et. al. 2005) claim to shape search: the properties of the surrounding environment and the goals of the observer. Encoding of this behavior must occur with a simple yet empirically supported method to garner support of the model development community. Promising avenues of research to address these open problems include representation of pre-attentive processing, incorporating peripheral detection, surmising threat from information, and extracting semantic meaning from physical surroundings (eg, hiding locations).

6. ACKNOWLEDGEMENT

This work would not have been possible without the support of the US Army Training and Doctrine Command Analysis Center (TRAC) and the Naval Postgraduate School. Significant experimental support was provided by Chris Darken, Patrick Jungkunz, Larry Shattuck, and Nita Miller. This acknowledgement does not imply approval of the results in this paper by these organizations or individuals. The results and opinions in this paper have been formulated and expressed solely by the authors.

7. REFERENCES

Cooke, K. J. “Modeling of Visual Search Performance”. British Aerospace Dynamics Group. 1983.

Darken, C. and Jones, B. "Computer Graphics-Based Target Detection for Synthetic Soldiers", Proceedings of Behavioral Research in Modeling and Simulation, 2007.

Doll, Theodore J. and Richard Home. “Lessons Learned in Developing and Validating Models of Visual Search and Target Acquisition.” Proceedings of the RTO SCI Workshop on Search and Target Acquisition, Utrecht, Netherlands, June 1999.

Eaton, Jeffrey P. “Analysis of the Effects of Dynamic Characteristic Dimension Calculations on FLIR Performance Prediction Models”. Naval Postgraduate School Master’s thesis, March 1997.

Endsley, Mica R., Leonard D. Holder, Bruce C. Leibrecht, Daniel J. Garland, Richard L. Wampler, and Michael D. Matthews. “Modeling and Measuring Situation Awareness in the Infantry Operational Environment”. U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report 1753, June 2000.

Glumm, Monica M., Kathy L. Kehring, and Timothy L. White. “Effects of Tactile, Visual, and Auditory Cues About Threat Location on Target Acquisition and Attention to Visual and Auditory Communications”. ARL-TR-3863, August 2006.

Glumm, Monica M., Kathy L. Kehring, and Timothy L. White. “Effects of Visual and Auditory Cues About Threat Location on Target Acquisition and Attention to Auditory Communications”. ARL-TR-3560, August 2005.

Hasting, Matthew D. “Combat Simulation of Individual Soldier Search in Urban Terrain.” Naval Postgraduate School Master’s thesis, June 2009.

Hoffman, Donald D. *Visual Intelligence*. W.W. Norton and Company, New York, New York, 1998.

Itti, L., and Koch, C. "A Saliency-based Search Mechanism for Overt and Covert Shifts of Visual Attention". *Vision Research*, 40, 1489-1506.

Johnson, J. "Analysis of Image Forming Systems". Proceedings of the Image Intensifier Symposium, pp. 249-273, U.S. Army Engineer Research and Development Lab, Ft. Belvoir, VA. (DTIC report AD 220 160), Oct. 6-7, 1958.

Jungkunz, Patrick. "Modeling Human Visual Perception for Target Detection in Military Simulations." Naval Postgraduate School Ph.D. dissertation, June 2009.

Mazz, John, Regina W. Kistner, and William T. Pibil. "Detection of Low-Contrast Moving Targets." ADP010547, June 1999.

Medina, John. *Brain Rules.* Pear Press, Seattle, 2008.

O'Kane, Barbara L., Irving Biederman, and Eric E. Cooper. "Modeling Parameters for Target Identification: A Critical Features Analysis." Proceedings of the IRIS Passive Sensors Symposium, Monterey, CA, March 1996.

Ratches, James A., Richard H. Vollmerhausen, and Ronald G. Driggers. "Target Acquisition Performance Modeling of Infrared Imaging Systems: Past, Present, and Future." IEEE Sensors Journal, Vol. 1, No. 1, June 2001.

Shattuck, Lawrence G., and Nita L. Miller. "Naturalistic Decision Making in Complex Systems: A Dynamic Model of Situated Cognition Combining Technological and Human Agents". Organizational Behavior: Special Issue on Naturalistic Decision Making in Organizations. Vol. 27, No. 7, 989-1009 2006.

Vaughan, Barry D.. "Soldier in the Loop Target Acquisition Performance Prediction Through 2001: Integration of Perceptual and Cognitive Models". ARL-TR-3833, July 2006.

Vollmerhausen, Richard H., Eddie Jacobs, and Ronald G. Driggers. "New Metric for Predicting Target Acquisition Performance." Optical Engineering, Vol. 43, No. 11, November 2004.

Wagner, D. H., Mylander, C. W., & Sanders, T. J. (1999). *Naval Operations Analysis, 3rd ed.* Naval Institute Press.

Washburn, A. R. *Search and Detection, 4th ed*. Institute for Operations Research and the Management Sciences, 2002.